AMENDMENT

In the Specification:

Please replace paragraph [68] as follows:

"**[68]** 3. Divide each dataset into approximately equal-sized virtual pieces, hereinafter referred to as data elements:

A. Choose data element size based on computational intensity of comparison algorithm (BLAST or Smith-Waterman), desired ratio of computation time to data transmitted, and available storage and RAM on slave CPUs.

B. The randomization of sequences can be done as in step 1, above, or within Prospector ™ itself. When done by Parabon's Prospector ™, if either the query or ~~sujeet~~ subject database set contains related sequences in a contiguous arrangement, randomize the order of sequences among the data elements by assigning each query or subject database sequence to a data element with the least size.

C. If individual database entries are larger than the desired data element size:

i. divide database entry into two smaller, overlapping pieces in two data elements, or

ii. choose to put large database entry into its own oversized data element.

D. If individual database entries within a data element are larger than the maximum desired datasize:

i. divide individual database entries within a data element into multiple entries of desired maximum length with overlap. Overlap of 50% ensures that no query/subject match less than the overlap length will be missed.

E. Strip all metadata from database entries. Context may be reconstructed with location information placed into data elements.

F. Pack data into efficient structure, e.g. 2 bits per nucleotide with appropriate encoding, 5 bits per amino acid residue with appropriate encoding, etc.

G. Create index for data and pack index and data into uncompressed data element structure.

H. Compress data into compressed data element structure with standard redundancy reduction data compression method, e.g. gzip, pkzip, etc."

B1